NPS-OR-95-001

NAVAL POSTGRADUATE SCHOOL

Monterey, California

ANALYSIS OF RECRUITING BONUS PAYMENTS

Harold J. Larson

February 1995

Approved for public release; distribution is unlimited.

Prepared for:
U.S. Army Recruiting Command
Fort Knox, KY 40121

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral T. A. Mercer
Superintendent

Harrison Shull
Provost

This report was prepared for and funded by U.S. Army Recruiting Command, Fort Knox, Kentucky, 40121.

This report was prepared by:

HAROLD J. LARSON
Professor of Operations Research

Reviewed by:

PETER PURDUE
Professor and Chairman
Department of Operations Research

Released by:

PAUL J. MARTO
Dean of Research

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
February 1995

3. REPORT TYPE AND DATES COVERED
Technical

4. TITLE AND SUBTITLE
Analysis of Recruiting Bonus Payments

5. FUNDING NUMBERS

6. AUTHOR(S)

Harold J. Larson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

NPS-OR-95-001

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Recruiting Command
Fort Knox, KY 40121-2726

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Army Recruiting Command (USAREC) has provided the number of enlistment contracts signed, by high school graduates in categories I through IIIa, for fiscal years 1988 through 1993. Enlistment bonuses are offered to attract these individuals to the Army, and to aid in channeling them into specific Military Occupational Specialties (MOSs). The bonus amounts offered, and the particular MOSs with which they were associated, varied over the period studied.

This study analyzed the numbers of contracts signed in 8 different MOS categories over these 6 fiscal years, seeking to identify the "best" enlistment bonus to offer for each of these MOSs. A simple linear spline model with one knot was used for each MOS; weighted least squares was used to estimate the parameters of the spline, including the location of the knot. This knot location may then be used to identify the enlistment bonus to offer for that MOS.

14. SUBJECT TERMS
Recruiting Bonus, Linear Splines, Least Squares

15. NUMBER OF PAGES
25

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

Analysis of Recruiting Bonus Payments

Harold J. Larson
Naval Postgraduate School

Abstract

The United States Army Recruiting Command (USAREC) has provided the number of enlistment contracts signed, by high school graduates in categories I through IIIa, for fiscal years 1988 through 1993. Enlistment bonuses are offered to attract these individuals to the Army, and to aid in channeling them into specific Military Occupational Specialties (MOS*s*). The bonus amounts offered, and the particular MOS*s* with which they were associated, varied over the period studied.

This study analyzed the numbers of contracts signed in 8 different MOS categories over these 6 fiscal years, seeking to identify the "best" enlistment bonus to offer for each of these MOS*s*. A simple linear spline model with one knot was used for each MOS; weighted least squares was used to estimate the parameters of the spline, including the location of the knot. This knot location may then be used to identify the enlistment bonus to offer for that MOS. The recommendations made are described in the following table.

Recommended four year bonus values, October 1990 dollars.

MOS	Amount	Comments
11X	$2915	Very strong recommendation
13B	$3145	Strong recommendation
13M	$2252	Strong recommendation
16S	****	Not appropriate model
19D	$3244	Weak recommendation
63B	$2729	Moderately strong recommendation
63T	****	Not appropriate model
94B	$1497	Weak recommendation

Introduction

Since the end of the draft in 1972, the Department of Defense has depended on volunteers to maintain necessary force levels. While there is no apparent shortage of persons willing to serve in the military, the increasing level of sophistication needed to employ modern weaponry brings its own demands on persons recruited to active duty. Those persons with greater skills and education have a wider range of employment possibilities open to them; the military thus is forced to compete with an increasing range of other options as it tries to recruit people with skills demanded by modern force structures.

The Department of Defense, and the Army in particular, employ aptitude and educational attainment in determining quality of enlistees. Persons who have a high school

diploma, and score 50 or higher on the Armed Forces Qualification Test (AFQT), are designated "high quality"; such recruits are sought because they are likely to complete their full term of enlistment and to be successful on the job.

Congress has authorized the use of a number of different incentives to help the military attract high quality persons into critical Military Operational Specialties (MOS*s*). Among the incentives offered is a cash bonus, paid to the recruit in installments over the agreed-to term of service. The bonus amounts offered, and the MOS categories to which they apply, are changed periodically; for the Army these decisions are made jointly by the Army Deputy Chief of Staff Personnel (DCSPER), the United States Army Recruiting Command (USAREC) and the Training and Doctrine Command (TRADOC). A committee representing these entities is occasionally called together to set the bonus amounts and the MOS categories covered for a coming period of time; this bonus information then is promulgated through the recruiting command by message, spelling out the MOS catgories covered, bonus amounts authorized, and applicable time limits, among others.

Over the six most recent fiscal years (1988–1993) a total of 28 different bonus periods occurred for Army enlistees. Within each of these periods the MOS categories which were offered cash bonuses, as well as the amounts offered, did not change. Any given MOS might have the same dollar value bonus authorized over several of these periods, but in general the bonus value shifted several times. For example, the four year enlistment bonus for MOS 11X (Infantry) had 7 different monetary values over these years (28 periods) from a low of $3,000 to a high of $8,000.

The enlistment bonus is intended to help channel high quality applicants into Army MOS categories in need of recruits. Presumably, the larger the bonus offered the greater the attraction to a prospective enlistee. With finite recruiting budgets, it is important to conserve resources; ideally, the amount of enlistment bonus offered should be (just) sufficient to attract the needed recruits for a given MOS. This "best" bonus amount to offer, for a given MOS, is not easily identifiable; it may be affected by a number of other variables, such as civilian opportunities available, offers made by other military services, bonuses paid for other MOS categories, and possibly many others.

This study takes a simple, rather gross approach to the investigation of the "best"

enlistment bonus to offer. It employs historical data on the numbers of contracts signed, and linear splines, to identify a critical bonus amount within the range of bonuses offered; this critical bonus amount is estimated from the data themselves and may, for some MOS*s*, be indicative of the "best" bonus to offer. This procedure is described in the following section.

Procedure employed

A linear spline with one knot is simply a straight line which has been bent at one location (the knot). Figure 1 uses a linear spline to represent a fictitious relationship between the amount of enlistment bonus offered (on the horizontal axis) and the resulting production of contracts signed per day (on the vertical axis). As pictured, this spline has a single knot at a bonus offer of \$2,800; bonus amounts greater than this have very little effect on the production of contracts, while smaller amounts rapidly reduce the contract production. For such a situation, \$2,800 would be attractive as the "best" bonus to offer.

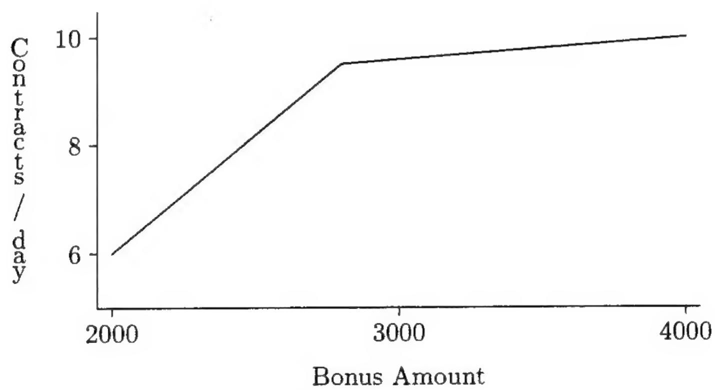


Figure 1. Linear spline

This picture and discussion are indicative of the analysis performed for this study. A recent paper [1] describes a simple procedure for using observed data, with least squares, to estimate a linear spline with one knot; the parameters of the spline, including the knot location itself, are all estimated from the data. The discussion in [1] explicitly treats only the simplest case, in which the variances of the dependent variable (contracts produced) are assumed equal; it does not explore statistical issues such as the apparent legitimacy of the knot located.

The current study uses historical data on the production of Army enlistment contracts

with the methodology of [1] to estimate a linear spline for certain selected MOS categories. This methodology is extended to use weighted least squares (appropriate for differing variances of the dependent variable) and to formally test certain hypotheses about parameters of the spline. An appendix discusses justifications for these extensions.

Data employed

USAREC provided descriptions of the bonus structure employed from fiscal year 1988 through 1993. As mentioned earlier, this time frame was partitioned by Army design (exhaustively) into 28 periods, within each of which the bonus structure was static. These time periods were quite variable in length, with the shortest covering 26 calendar days and the longest 171 calendar days. Figure 2 presents a picture of the 28 periods into which these 6 fiscal years were partitioned, and the four year bonus amounts available to an 11X enlistee; the 28 ticks on the horizontal axis (every fourth one is numbered) locate the midpoints of the 28 periods; the horizontal axis spans the time period October 1, 1987, through September 30, 1993. Period 6 was 171 days long, while period 26 was 26 days long.

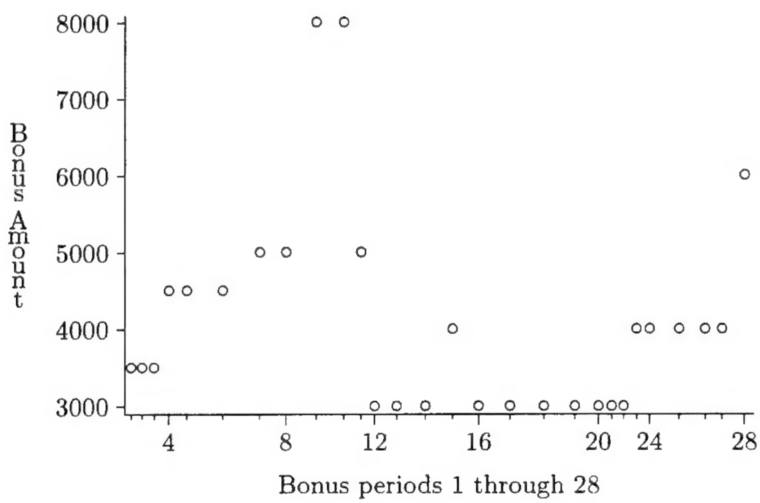


Figure 2. Bonus for 11X and bonus periods

USAREC also provided a data set describing the 375,233 high quality contracts signed over this time period; this data set includes such things as the MOS contracted for, date of contract signing, bonuses available, age, education, gender, and term of service, among others. It is important to recognize that the data analysed concerns contracts signed, as opposed to enlistees actually accessed into the Army; the great majority of those persons

signing contracts to enlist take part in the Delayed Entry Program, allowing them time to reconsider their choice to enlist. Some of those signing enlistment contracts chose not to honor their commitment during this period.

Not all MOS categories are authorized to use enlistment bonuses; the first step in this study was to identify a number of MOS*s* which satisfied two requirements:

a. The four year enlistment bonus offer varied over the total time span.

b. The total number of contracts signed was fairly large.

These requirements led to the MOS*s* listed in Table 1.

Table 1. MOSs selected for analysis

MOS	Identification	Number of contracts
11X	Infantry	56,679
13B	Cannon Crewmember	10,405
13M	Multiple Launch Rocket System Crewmember	3,054
16S	Man Portable Air Defense System Crewmember	2,376
19D	Cavalry Scout	6,680
63B	Light Wheel Vehicle Mechanic	7,620
63T	M2-3 Bradley Fighting Vehicle System Mechanic	2,643
94B	Food Service Specialist	5,848

In the analyses performed, the total number of contracts signed in the specific MOS was used, not just those for four years, nor just those whose contract included the enlistment bonus. This approach models the gross effect of four year bonus amounts on total contracts signed in the MOS.

A number of different models have been investigated. The numbers of contracts signed can be modelled strictly for the 28 periods already mentioned (grouping the data into just these 28 periods). To possibly allow more detailed investigation, the numbers of contracts signed can also be modelled on a (roughly) monthly basis, over the 72 months spanned by the 6 fiscal years. Analyses have been done both ways, producing essentially the same results; the results for monthly grouping of the data will be explicitly addressed. The main quantity of interest is the location of the knot, the critical bonus amount at which the rate of contract production seems to change; the slopes on either side of the knot (at least in terms of their signs) are also of interest. The major independent variable is the bonus offered;

since the data spans a relatively long period of time, the consumer price index (using values reported in [2]) has been used to express all dollar amounts in constant October, 1990 dollars, roughly the midpoint of the full time interval spanned. Unemployment rates for youths aged 16 to 19 are available from [3]; these were also investigated for use in modelling the production of contracts for these selected MOS categories.

Model used for the data

The observed time span covers 72 calendar months; the boundaries between the periods where a change occurred in the bonus amount offered did not always fall at the end of the month. For each individual MOS studied (listed in Table 1), the numbers of enlistment contracts were subdivided into calendar months when the four year bonus offer allowed; in those cases in which the bonus offer changed, the boundary of what was called a month coincided with the date of the change in the bonus. Within each of these 72 defined periods, the contract dates for the given MOS were counted. For MOS 11X, the numbers of contract dates (within the defined periods) varied from 17 to 28, with a median of 23 days. Specialty 16S, the MOS with the smallest number of contracts, presented a problem with this approach; the calendar months of August and September, 1991, contained no contract dates. Thus the monthly sample size used for MOS 16S is 70, not 72 as for all the other MOS categories analysed. Once the numbers of contract dates per period was established (for each MOS), the number of contracts signed for each such date was also determined; the total number of contracts in the month divided by the number of contract days was used as the dependent variable in this study.

For example, 694 contracts for MOS 11X were signed in October, 1987, which contained 24 observed contract dates; the mean number of contracts per (contract) day then is $694/24 = 28.92$ for this MOS for this month. The standard deviation of the number of contracts per day has also been evaluated: this value is 14.71 for MOS 11X for October, 1987. The estimated standard error of the mean number of contracts for MOS 11X for October, 1987, is $14.71/\sqrt{24} = 3.003$. These estimated standard errors were used in applying weighted regression to estimate the linear spline. Please note that the standard deviation of the number of contracts per day is 0 whenever the number of contracts written is the same constant for each day; since the appropriate weights are the reciprocals of the standard

deviations, these weights will be undefined for any month with a zero standard deviation. This situation does occur in these data; MOS 16S has 6 months for which the standard deviation is 0, while 16D has 3 and 63T has 1. This was handled by setting the standard deviation for any month in which it was 0 equal to the smallest monthly standard deviation for that MOS among the remaining months.

For each monthly period the four year enlistment bonus is known; as mentioned earlier, these values were adjusted by the consumer price index and expressed in constant 1990 dollars. Figure 3 plots the resulting mean number of contracts per month, for MOS 11X for the 72 months, versus the (cpi adjusted) enlistment bonus offered. It is interesting to note that the largest daily rate of contract production occurs with the lower, not the higher, enlistment bonuses. (It is universally true for all 8 MOS categories investigated that the highest observed rate of contracts/day occurs with the four year enlistment bonus smaller than the largest used; see the figures in Appendix A). This behavior could be anticipated: since the bonus levels are set to channel potential enlistees into specific MOS categories, those periods with high bonus offers should coincide with conditions where the natural choice tendencies would not produce the number of contracts needed. If the bonus offer were lower the contract production realized for these periods would probably have been lower as well.

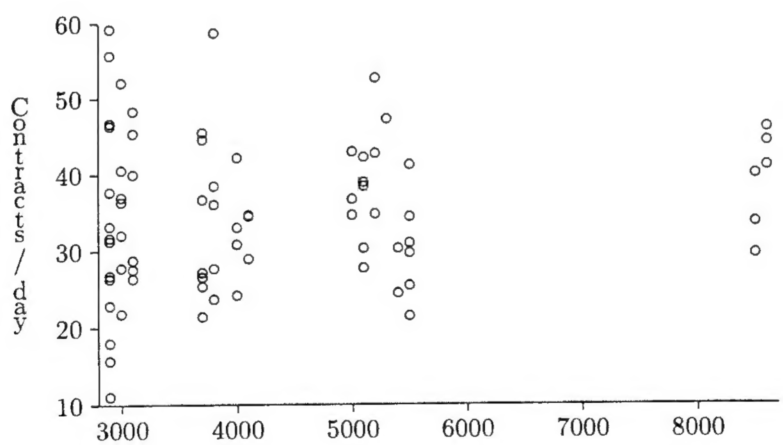


Figure 3. MOS 11X. contracts/day versus bonus

A linear spline with one unknown knot location has been fit to these monthly data for

MOS 11X. The assumed model can be written

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + e_i \qquad \text{for } x_i \le a \\ &= \beta_0 + \beta_1 a + \beta_2 (x_i - a) + e_i \qquad \text{for } x_i \ge a, \end{aligned} \tag{1}$$

where y_i is the average number of 11X contracts per day for month i, x_i is the (cpi adjusted) four year enlistment bonus offered, β_0 is the y-intercept, a is the unknown knot location, β_1, β_2 are the slopes to the left and right, respectively, of the knot, and e_i is the residual for month i. The e_i values are assumed to have mean 0 and different variances; these variances may differ for two possible reasons:

- The number of contract days varies from month to month.
- The variability in contracts per day does not seem to be constant across the 72 months.

As mentioned earlier, the standard errors of the mean contracts per day have been evaluated; the reciprocals of these values have been used as weights and weighted least squares has been applied to estimate the unknown coefficients in the model. The appendix has a discussion of extending [1] for weighted least squares. For MOS 11X, the resulting estimates are $\hat{\beta}_0 = -935.69, \hat{\beta}_1 = 0.33176, \hat{\beta}_2 = 0.00021$ and the estimated knot location is $\hat{a} = 2915.45$. This estimated linear spline is plotted over the observed data in Figure 4.

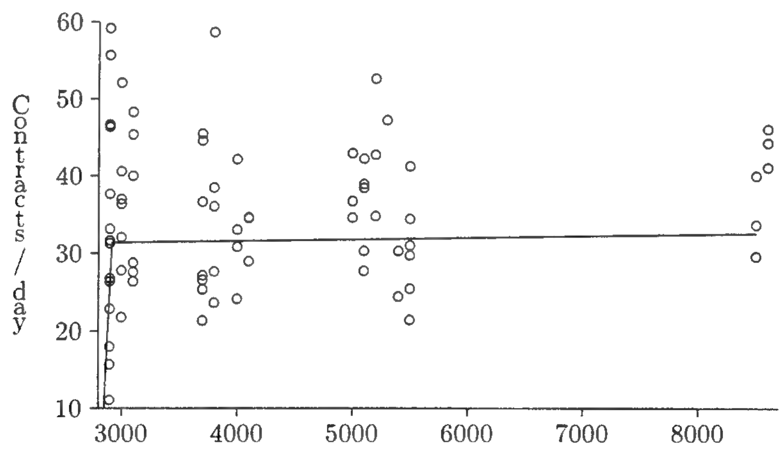


Figure 4. MOS 11X, best linear spline, weighted

The estimated standard deviations of the $\hat{\beta}_i$ values are 190.30, 0.06556, 0.00072, respectively (computed in the usual way). The ratios of $\hat{\beta}_1, \hat{\beta}_2$ to their standard errors are 5.06, 0.29, respectively; if these behave like t random variables the data are consistent with the

slope to the left of the knot being positive, while the slope to the right of the knot is 0. As discussed in Appendix B, an F-ratio with 2 and 68 degrees of freedom seems appropriate for judging whether the located knot is "real". (Does the observed data seem consistent with having no knot, within sampling error?) For MOS 11X, the value of this statistic is 12.68, with p-value .00002; the observed data are much better fit using a linear spline with one knot than by simply employing a straight line.

This same analysis has been performed as well for the other seven MOS categories listed in Table 1. The resulting knot locations, pairs of slopes, their t-values, and the significance of the F ratio described above are listed in Table 2. Note that MOS 19D and MOS 94B both have quite high p-values for the F-ratio, indicating that the spline with one knot does not fit the observed data any better than a straight line. For the other MOS categories, the spline does seem to perform better; MOS 16S and MOS 63T both have significant negative slopes to the left of the knot with this model. These are the two categories with the smallest number of contracts produced. (See the conclusions for more comments on these two.) This could be caused by the model not including one or more independent variables which are important for modelling their behavior. The original data values, and the estimated spline, are pictured in Appendix A for these additional MOS categories, in the same manner as Figure 4 for MOS 11X.

Table 2. MOSs selected for analysis, bonus as independent variable.

MOS	Knot value	Left slope (t value)		Right slope (t value)		p-value for F ratio
11X	2915.45	0.33176	(5.060)	0.00021	(0.292)	0.000
13B	3090.14	0.01381	(2.625)	−0.00002	(−0.035)	0.061
13M	2252.25	−0.00162	(−1.723)	0.00081	(5.400)	0.056
16S	967.12	−1.31634	(−2.581)	0.00012	(1.714)	0.042
19D	2741.23	0.00059	(2.565)	−0.00032	(−0.344)	0.706
63B	2729.26	0.00041	(2.158)	0.00590	(2.479)	0.088
63T	2169.20	−0.00499	(−2.458)	0.00030	(1.667)	0.046
94B	1485.15	−0.00040	(−1.212)	0.00016	(1.067)	0.432

As mentioned earlier, the monthly unemployment rates for youths aged 16 through 19 are available from [3]. These values have been used as well, as an additional independent variable in modelling the production of contracts for the 8 MOS categories. The model

now becomes

$$y_i = \beta_0 + \gamma u_i + \beta_1 x_i + e_i \qquad \text{for } x_i \le a$$

$$= \beta_0 + \beta_1 a + \gamma u_i + \beta_2(x_i - a) + e_i \qquad \text{for } x_i \ge a,$$

where all common terms have the same definition as in equation (1); u_i is the youth unemployment rate for month i and we have one additional parameter (γ) to estimate. Again weighted least squares was employed in estimating the unknown coefficients in this model. For MOS 11X, the resulting coefficient estimates are $\hat{\beta}_0 = -939.141, \hat{\gamma} = 0.55968, \hat{\beta}_1 = 0.32967, \hat{\beta}_2 = 0.00022$; their estimated standard errors are, respectively 189.418, 0.43555, 0.065267, 0.000716. The estimated knot location is 2915.45, the same value as before. The estimate for γ is positive as would be expected (since an increase in the unemployment rate should make employment in the military more attractive), but is not significant. The F ratio for testing whether the knot seems real is 12.63, again highly significant. Table 3 presents the estimates of the knot location, the three coefficients γ, β_1, β_2, each followed by its corresponding t value (coefficient divided by standard error) and the p-value for the F ratio testing whether the linear spline fits the data any better than a simple straight line.

Table 3. Unemployment and bonus as independent variables.

MOS	Knot value	$\hat{\gamma}$	(t value)	Left slope	(t value)	Right slope	(t value)	F p-value
11X	2915.45	0.55968	(1.28)	0.32967	(5.05)	0.00022	(0.31)	0.000
13B	3272.75	0.10858	(0.90)	0.00653	(2.60)	−0.00009	(−0.15)	0.094
13M	2252.25	−0.00835	(−0.24)	−0.00163	(−1.73)	0.00082	(5.36)	0.057
16S	967.12	−0.03325	(−0.88)	−1.37135	(−2.66)	0.00011	(1.63)	0.034
19D	3243.74	0.18684	(1.71)	0.00034	(1.83)	0.00159	(0.83)	0.823
63B	2729.26	−0.12270	(−1.29)	0.00040	(2.04)	0.00587	(2.48)	0.087
63T	2169.20	0.02220	(0.63)	−0.00496	(−2.43)	0.00030	(1.62)	0.049
94B	1485.15	0.07440	(1.36)	−0.00043	(−1.32)	0.00015	(0.98)	0.409

As with the earlier model which employed only the four year enlistment bonus as an independent variable, the F ratio p-values are high for 19D and 94B; the spline does not seem to fit the observed data much better than a simple multiple regression for these two categories. For the other six MOSs, these p-values are roughly the same as they were for the simpler model in equation (1) and, except for MOS 13B, the knot locations are identical (the shift of \$182 for 13B does not seem particularly large). The slopes to the left and right of the knot are also quite similar to those estimated for equation (1). The coefficient

γ is not significant for any of these MOS*s*, saying the simpler model with just the four year enlistment bonus as an independent variable performs about as well for any of them. Thus, the unemployment data was not used in any further analysis.

Variability of knot location

The standard least squares models provide easy estimates of the standard deviations of the estimated parameters, when the mean is a linear function of the unknown parameters. The mean function for this spline model is not a linear function of the unknown parameters $\beta_0, \beta_1, \beta_2, a$, so these standard formulas are not directly applicable in estimating the variability of the estimates. However, if as above the residuals are assumed to be normally distributed, Fisher's information matrix can be used to evaluate the asymptotic standard deviations of the parameter estimates; this leads to the values used earlier to compute the t values given for the slope coefficients. The asymptotic standard deviations for the knot locations ($\hat{a}$ values) have also been evaluated and are given in Appendix B. In this section, we discuss a sampling-based method of investigating the variability of the knot locations.

The "bootstrap" procedure provides a simple way to estimate the variability of an unknown parameter. This procedure calls for repeatedly sampling at random, with replacement, from the observed data; for each such sample the parameter estimate is evaluated. The variability of these generated estimates should then be representative of the "true" variability of the estimator employed.

This approach has been employed to investigate the variability of $\hat{a}$, the estimated knot location, for each of the MOS categories. Specifically, the 72 months of data for MOS 11X contain varying numbers of contract days, and each contract day produced varying numbers of contracts signed. There were 24 contract days in October, 1987, 22 in November 1987, and so on, with 22 contract days in September, 1993. To investigate the variability of $\hat{a}$ for 11X, 100 independent estimates were generated: this was done by randomly selecting 24 of the numbers observed for October, 1987, with replacement, together with 22 numbers selected from the observations for November, 1987, with replacement, and so on, 100 times. For each generated sample the same procedure was used to estimate the least squares spline. This then generated 100 values for $\hat{a}$: this simulation of 100 samples was done in turn for each of the MOS*s*. Table 4 presents the average knot values, the standard deviations, the

medians, as well as the 6th and 95th ranked values of the 100 (these last two values define an interval which contained 90 of the 100 observed values).

Table 4. Summary statistics for simulated knot values.

MOS	Mean	Standard Deviation	Median	6^{th} quantile	95^{th} quantile
11X	2915.18	3.28	2915.45	2908.14	2918.29
13B	3468.61	672.98	3145.36	2901.35	4569.32
13M	2357.99	413.17	2252.25	1917.55	3371.87
16S	1092.61	601.97	967.12	967.12	1470.59
19D	3145.62	504.56	3243.74	2085.51	3751.34
63B	2615.41	145.60	2729.26	2310.54	2729.26
63T	2324.82	370.33	2169.20	2143.62	3234.75
94B	2295.66	1065.29	1497.01	1485.15	4338.39

As is not surprising, the median of the observed values is typically quite close to (if not identical with) the actual estimated knot value. The simulated knot values for 11X varied hardly at all; the width of the interval containing 90 of the simulated results is \$14.15. Those for 94B varied the most; the width of the interval containing 90 of the simulated values is \$2853.24 (recall that 94B is one of those for which the linear spline is not significant). Note that these generated sampling distributions are quite skewed for several of these MOS categories; this can be observed by comparing the mean with the median, or perhaps more clearly by observing the location of the median versus the 6^{th} and 95^{th} quantiles. For example, the median knot for 13B is 3145.36, while the two extreme quantiles are 2901.35 and 4569.32, respectively. Thus this distribution is quite severely skewed to the right (as are several of the others, with a few skewed to the left).

Conclusions

This crude model gives clear indications for MOS categories which require large numbers of recruits; it does not seem to work well for the MOS categories 16S and 63T, which had the two smallest contract productions. MOS 16S has a number of months with a small number of contract days (there are two months in which there were none); in addition, many of these contract days have exactly one contract signed. MOS 63T has relatively more recruiting days per month, but also frequently has just one contract signed on a given day. This type of behavior may be indicative of manipulation of the system through

which the recruiting counsellors are given the opportunities which they may then present to prospective contract signers. A counsellor may call the USAREC Recruiting Operations Command and specifically ask for a given MOS assignment, which is not on his system at the given time, to "lock in" a contract for a given person; permission may allow the counsellor to sign a contract for an MOS (and given accession date) which is otherwise not generally available. Such behavior would likely lead to many recruiting days with a single contract signed; cases in which it is common to have one contract per day (and many days with no contracts) should be modelled in a different way.

Table 5. Recommended four year bonus values, October 1990 dollars.

MOS	Amount	Comments
11X	$2915	Very strong recommendation
13B	$3145	Strong recommendation
13M	$2252	Strong recommendation
16S	****	Not appropriate model
19D	$3244	Weak recommendation
63B	$2729	Moderately strong recommendation
63T	****	Not appropriate model
94B	$1497	Weak recommendation

The linear spline model fits the data for 19D and 94B no better than a simple straight line. While the spline does identify a knot, a monetary value at which the rate of production of contracts changes, the indication is not strong that this value has any real meaning for these two cases. For the remaining four MOS categories (11X, 13B, 13M, and 63B) the data presents moderate to strong indicators of the "best" bonus to offer (within the range of bonuses used). The bonus recommended is the median of the boot strap samples discussed earlier, rounded to the nearest (October 1990) dollar. The bonuses actually used, of course, would logically be multiples of $500 or $1000, as have been used historically. These conclusions are summarized in Table 5.

References

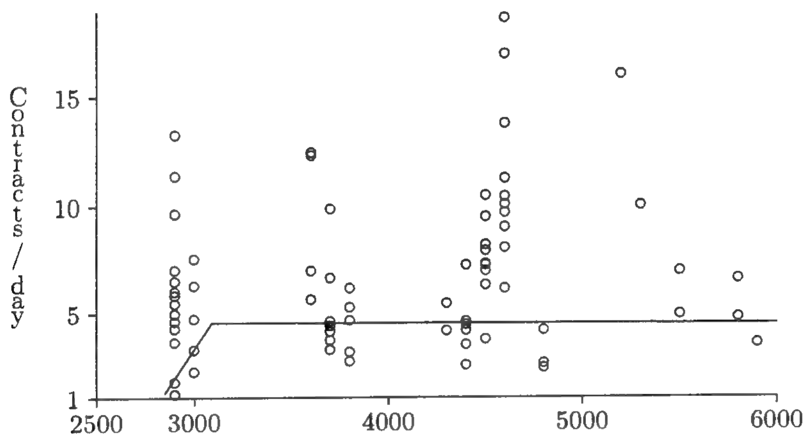
[1] Larson, Harold J., "Least squares estimation of linear splines with unknown knot locations", *Computational Statistics & Data Analysis*, Vol. 13, pp 1–8, January, 1992.

[2] Bureau of Labor Statistics, US Department of Labor, "Consumer Price Index Detailed Report", September 1993.

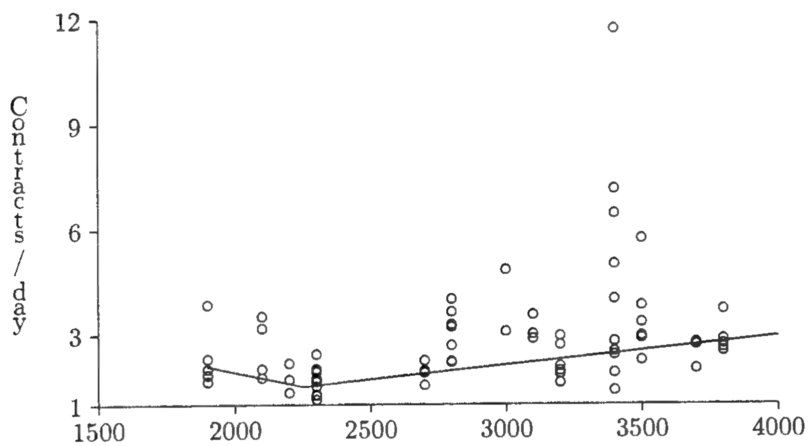
[3] Council of Economic Advisers, "Economic Indicators", September 1993.

Appendix A

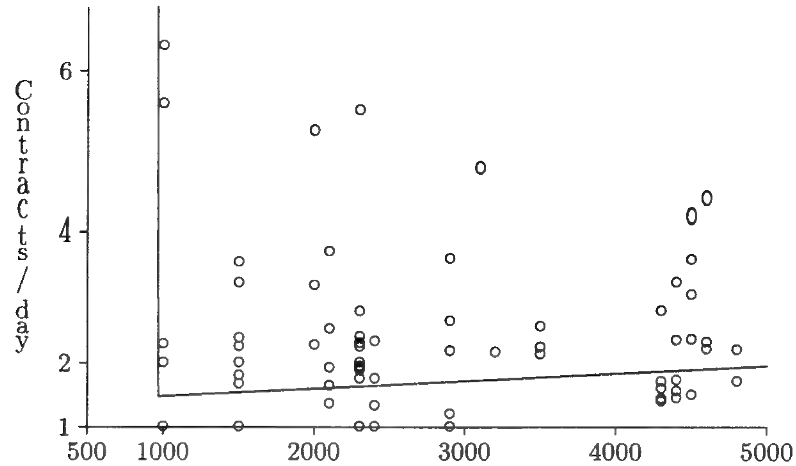
Presented here are pictures of the observed data for MOS categories 13B, 13M, 16S, 19D, 63B, 63T and 94B. Please note carefully the changes of scale used for both the horizontal and vertical axes. The mean numbers of contracts signed per contract day, versus the (cpi adjusted) four year bonus amount are represented by small circles; the best fitting (weighted least squares) spline with one knot is overlaid on the observed data for each MOS presented.



Four year enlistment bonus (October 1990 dollars)
Figure 1a. MOS 13B, best linear spline, weighted

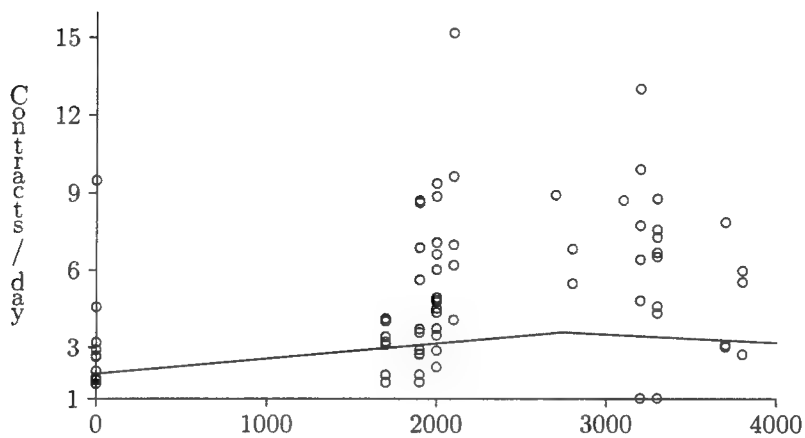


Four year enlistment bonus (October 1990 dollars)
Figure 2a. MOS 13M, best linear spline, weighted



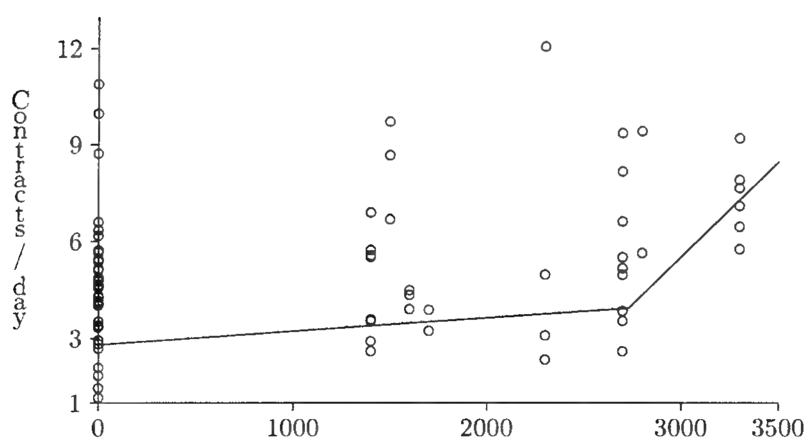
Four year enlistment bonus (October 1990 dollars)

Figure 3a. MOS 16S, best linear spline, weighted



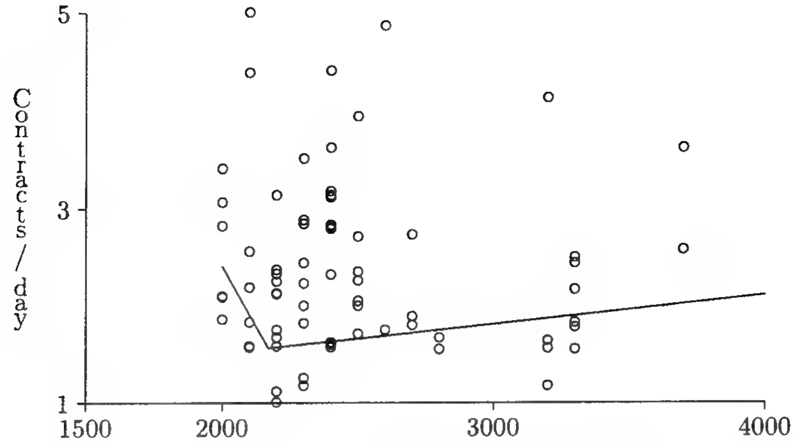
Four year enlistment bonus (October 1990 dollars)

Figure 4a. MOS 19D. best linear spline. weighted



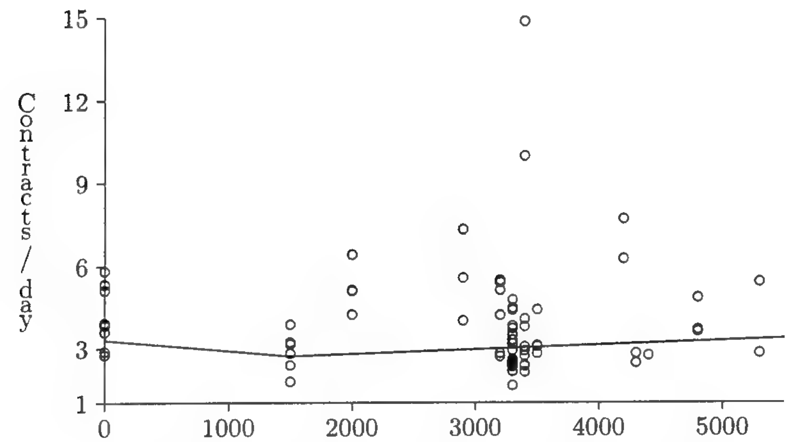
Four year enlistment bonus (October 1990 dollars)

Figure 5a. MOS 63B. best linear spline. weighted



Four year enlistment bonus (October 1990 dollars)

Figure 6a. MOS 63T, best linear spline, weighted



Four year enlistment bonus (October 1990 dollars)

Figure 7a. MOS 94B, best linear spline, weighted

Appendix B. Some technical details

The basic methodology employed for this study is described in [1], with the difference that β_0 here represents the y-intercept, not the height of the spline at the estimated knot. As already noted, there are reasons to believe that the dependent variables used in these analyses may have different variances (driven by the facts that the variances of contracts signed, per month for a given MOS, seem to differ, and that the numbers of contract days observed were not the same for all months). This calls for a straightforward adjustment to the procedure used, as can be sketched in as follows for the simple model using only enlistment bonus as an independent variable; exactly the same argument is appropriate and goes through with no difficulty for a model with two or more independent variables.

Equation (1) states the model as

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_i + e_i \qquad \text{for } x_i \le a \\ &= \beta_0 + \beta_1 a + \beta_2 (x_i - a) + e_i \qquad \text{for } x_i \ge a; \end{aligned} \tag{2}$$

the residual error terms e_i are assumed to be uncorrelated with mean 0 and $\mathrm{Var}[e_i] = w_i^2\sigma^2$. As is well known, the best linear unbiased estimates then are produced by dividing this equation by w_i giving

$$\begin{aligned} \frac{y_i}{w_i} &= \beta_0 \frac{1}{w_i} + \beta_1 \frac{x_i}{w_i} + \frac{e_i}{w_i} \qquad \text{for } x_i \le a \\ &= \beta_0 \frac{1}{w_i} + \beta_1 \frac{a}{w_i} + \beta_2 \frac{(x_i - a)}{w_i} + \frac{e_i}{w_i} \qquad \text{for } x_i \ge a, \end{aligned} \tag{3}$$

an equation whose residuals e_i/w_i then have constant variances; simple least squares is then applied to this transformed model. This weighting of values has no real effect on the procedure described in [1] and implemented here, in terms of estimating the unknown coefficients, since the same minimizing argument is totally appropriate. These weights have been estimated from the observed data by the reciprocals of the standard errors of the mean number of contracts per month, as mentioned.

There are a number of more or less obvious results if one assumes normality (as well as independence and constant variance) of the residuals e_i/w_i. If Y is a random variable and $\mathrm{E}[Y \mid x]$ is assumed to follow the linear spline as given in equation (3), then the least squares estimates given in [1] and used here, including $\hat{a}$, are identical with the maximum likelihood estimates. This follows directly from the structure of the likelihood function:

maximizing the likelihood means minimizing the exponent of the normal joint probability density function. Minimizing this exponent is identical with the least squares procedure employed.

If the knot location a were assumed known (so there is no need to estimate its value), all of the classical linear statistical model theory is directly applicable. In particular, the estimators of the parameters of the spline (y-intercept and slopes) then are normal random variables, the residual sum of squares over σ^2 is a χ^2 random variable and this residual sum of squares is independent of each of the spline coefficient estimators. It then follows that Student's t distribution can be used for testing the significance of any single parameter in the spline (as has been done in judging the lack of significance of the unemployment variable); the F distribution is available for testing whether two or more coefficients of the spline have specified values.

For any given case with $x_i \leq a < x_{i+1}$ the (weighted) model given in equation (3) can be expressed with usual matrix notation as (also discussed in [1])

$$Y = X\beta + e = (M + aV)\beta + e.$$

That is, the vector β contains the y-intercept and the two slopes and the knot location a is involved in partitioning the X matrix into values to the left and right of the knot. For this model (assuming normality), Fisher's information matrix gives the variance-covariance matrix of $\hat{\beta}$ to be $\sigma^2(X^TX)^{-1}$, the same as the result for simple multiple regression. This is the basis for the estimated variances of the $\hat{\beta}_i$ values used in this report. The information matrix gives the asymptotic variance of $\hat{a}$ to be $\sigma^2(\beta^TV^TV\beta)^{-1}$. This in turn translates into the estimator

$$s_{\hat{a}}^2 = s^2/(\hat{\beta}_1 - \hat{\beta}_2)^2 \sum 1/w_i^2,$$

where s^2 is the residual mean square and the sum in the denominator is over those indices for which $x_i > \hat{a}$. These estimates have been evaluated for the MOS categories discussed and are presented in Table 1b, along with the estimated standard deviations produced by the bootstrap sampling procedure mentioned above. It is interesting how close together the two are for some cases, and how disparate they are for others. This disparity might be expected for the cases where the model does not seem appropriate; otherwise, such

differences can be caused by the fact that the asymptotic value (essentially assuming an infinite sample) is being compared to one based on 100 replications, as well as the rationale of the bootstrap itself.

Table 1b. Asymptotic standard errors versus bootstrap values.

MOS	Asymptotic s.e.	Bootstrap s.d.
11X	3.17	3.28
13B	25.77	672.98
13M	36.85	413.17
16S	.0603	601.97
19D	523.49	504.56
63B	202.41	145.60
63T	15.83	370.33
94B	243.49	1065.29

The use of the t distribution for judging significance of individual coefficients in this model is *ad hoc* but seems reasonable because of its appropriateness if the knot location were assumed known. An F test has been employed to judge whether or not the linear spline fits better than a simple straight line. If the knot is not "real" then $\beta_1 = \beta_2$, the slopes on the two sides of the knot are equal (and the knot has disappeared). This reasoning might superficially lead one to use the standard linear model t test of the hypothesis that $\beta_1 = \beta_2$ (or its square which has the F distribution with one numerator degree of freedom) to compute the p-value in testing whether the knot is "real". However, in proceeding from a model with no knot to one which includes a single knot (at estimated location $\hat{a}$) the number of coefficients to estimate increases by 2, not 1, since both a and another slope parameter must be estimated. The p-values quoted from the F distribution are based on this logic.

Specifically, let S_0 be the residual sum of squares using an estimated knot and let S_1 be the residual sum of squares if no knot is used: the difference $S_1 - S_0$ is treated as a constant times a χ^2 random variable with 2 degrees of freedom. The ratio $k(S_1 - S_0)/2S_0$ is used in getting these p-values; k is the number of degrees of freedom for S_0 (which is $72 - 4 = 68$ for 72 months with no unemployment variable and is $72 - 5 = 67$ for 72 months using the unemployment variable). The tail area to the right of this ratio, using the $F(2, k)$

distribution, is the quoted p-value. This ratio is the same as the square of the t random variable used to test that $\beta_1 = \beta_2$; in a certain sense this calculated p-value is conservative. The computed F statistic is .5 times the square of this t statistic; the area under the $F(2, k)$ distribution to the right of this F statistic (the p-value) is always greater than the area to the right of this t^2 value under the $F(1, k)$ distribution.

INITIAL DISTRIBUTION LIST

1. Research Office (Code 08) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

2. Dudley Knox Library (Code 52) 2
   Naval Postgraduate School
   Monterey, CA 93943-5002

3. Defense Technical Information Center 2
   Cameron Station
   Alexandria, VA 22314

4. Department of Operations Research (Code OR/Bi) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

5. Prof. Harold J. Larson (Code OR/La) 3
   Naval Postgraduate School
   Monterey, CA 93943-5000

6. Commander, USAREC 3
   Attn: RCPAE-RC (Maj. Poikonen)
   Fort Knox, KY 40121-2726